DOCUMENT RESUME

ED 179 930 CS 005 154

AUTHOR Gastr.ght, Joseph F.
TITLE Summer Loss in Reading Achievement: A Comparison of Fall to Spring Reading Achievement Gains with Fall to Fall Reading Achievement Gains.
PUB DATE Sep 79
NOTE 37p.: Paper presented at the Annual Meeting of the Evaluation Network (Cincinnati, OH, September 24-29, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Compensatory Education Programs: Elementary Education: Longitudinal Studies: *Program Effectiveness: *Reading Achievement: *Reading Improvement: *Retention: *Time Factors (Learning)
IDENTIFIERS *Elementary Secondary Education Act Title I: *Summer Vacations

ABSTRACT

The patterns of reading achievement growth for Title I children in Cincinnati, Ohio, schools were examined to see whether gains in reading achievement by these children were lost over the summer months. The subjects were 295 children who remained in the program from second to fourth grade, and 175 children who remained in the program from third to fifth grade. Reading achievement loss was measured in absolute terms (lower scores in fall than in spring) and in relative terms (smaller than expected gains between spring and fall). Loss in absolute reading achievement occurred in only one of four summer periods: and this absolute loss was related to changes in the achievement test levels administered during the spring and fall. When compared to national norms, relative losses in reading achievement were found in three of the four summer periods. During the two-year period, a significantly higher growth in reading achievement occurred for the Title I children than for other Cincinnati elementary school children. The summer gains and losses in relative achievement were not predictable from year to year for individual children. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED179930

SUMMER LOSS IN READING ACHIEVEMENT: A COMPARISON OF FALL TO SPRING READING ACHIEVEMENT GAINS WITH FALL TO FALL READING ACHIEVEMENT GAINS

by

Joseph F. Gastright
Evaluation Branch
Cincinnati Public Schools

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Joseph F. Gastright

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at annual meeting of the
Evaluation Network
Cincinnati, Ohio

1979

CS005154

Summer Loss in Reading Achievement: A Comparison of Fall to Spring Reading Achievement Gains with Fall to Fall Reading Achievement Gains

Several recent studies have highlighted the fact that Title I students achieve above expectation during the regular school year, but lose much of the achievement during the summer months. This study investigates the pattern of reading achievement gains made by two cohorts of Title I students tested five times in two years as part of the evaluation of the Growth in Academic Performance component of the Cincinnati ESEA Title I program. The study will address the following questions:

1. Do Title I students show absolute losses in reading achievement over the summer period?
2. Do Title I students show relative losses in reading achievement over the summer period when compared to the national norm group?
3. Do Title I students show relative losses in reading achievement over the two year period when compared to other children in the Cincinnati public schools?
4. Is the pattern of achievement gain made by Title I children consistent over two successive summer periods?

Literature Review

Thomas and Pelavin (1976) analyzed statewide aggregations of Title I data collected between 1970-71 and 1973-74. They noticed that the reported school year gains were consistently at or above the goal of 1.0 month achievement per month of program, but that in the successive grades there was little evidence of cumulative achievement growth. They recommended that future evaluation studies be based on multi-year records of achievement growth.

Pelavin and David (1977) analyzed several sets of data and concluded that the achievement rates differ dramatically when different time periods are used for the evaluation of compensatory education programs. They further stated that summer loss was consistent across different tests, grades, subjects and programs, and concluded that the evidence was sufficient to support the generalizability of the findings. DeVito and Long (1977) and Stenner (1977) corroborated the existence of summer loss using statewide assessment and local school district tests, respectively. DeVito and Long also reported substantially larger decreases when testing levels were crossed.

Holthouse, Stofflet and Tokor (1976) on the other hand, reported significant growth in reading across the summer months for students at several grade levels irrespective of the race or socioeconomic status of the students involved. The authors, however, used a technique which "renormed" the tests, and their results are not directly comparable to those reported in the other studies.

Measurement of Summer Loss. A distinction should be made between an absolute loss in reading achievement and a relative loss in reading achievement. Absolute summer losses in reading achievement occur when a group of children get a lower raw score on the fall administration of the same test they took in spring, or when they obtain a lower standard score on the fall administration of an equivalent test then the one they took in spring. In most cases, absolute loss in achievement has been reported in standard score units. Relative losses in reading achievement occur when a group of children make smaller than expected gains between the spring and fall administration of either the same test or a test equated to the original test. The relative losses represent a comparison of the

observed gain with the gain found in some comparison group. In most cases, relative losses in achievement have been reported in grade equivalents, percentile ranks or normal curve equivalents. The latter measure is simply a normalized standard score based on the percentile rank norms of a standardized test.

Absolute Measures of Reading Loss. The expanded standard score scale is the metric used by all test publishers to provide an equal interval scale which ties together the separate test levels and forms which make up a test battery. In general, the standard score scale is produced by applying the Thurstone scaling procedure to the empirically derived raw scores obtained on each form of the test. The same students are given two different forms or levels of the test and the raw score to standard score scales are then created by making the assumption that students will attain the same percentile rank on each test. These equipercentile norms are then extended up and down from some anchor test to create the expanded standard score scale. In theory, the standard score scale provides a uniform measure of the underlying achievement tested by the test battery. The process of establishing standard score scales across levels of a test requires fitting of the best smooth curve across the median values at adjacent norm points. Ozenne (1978) hypothesized that the size and composition of the norm group affects the amount of smoothing and curve fitting that must be done. He points out that the curve fitting methods are likely to introduce the greatest error at the extremes of the distribution where Title I students are located. Both Thomas and Pelavin (1976) and Pelavin and David (1977) report absolute losses

in standard scores across the summer months. In both cases, however, the standard scores were "recreated" from the mean grade equivalent scores, and the reported losses do not necessarily represent actual declines in the standard scores.

In any event, standard score changes across the summer period represent the closest approximation available on standardized tests to an absolute measure of summer loss. Raw score declines would be a more direct measure of summer loss. However, raw score declines could only be used in cases where the same form and level of the test were given in spring and later in fall.

Relative Measures of Reading Loss. The first studies based on Title I evaluation data all used the grade equivalent gains which were required by the existing Title I evaluation plans. Every study utilizing grade equivalent gains has pointed out the inherent problems in the grade equivalent scale. Thomas and Pelavin (1977) state that grade equivalent gains are not equivalent across tests, and that they are not equivalent from grade to grade or person to person within the same test. They also pointed out that all grade equivalent scores are based on equal month interpolation of the grade equivalents between the annual norm points. Only two studies are reported which tested the accuracy of these artificial norm points. Beggs and Hieronymus (1966) reported that the actual achievement gains made from January to April were greater than those implied by the grade equivalent scale of the Iowa Test of Basic Skills. They also reported that empirical summer losses occurred in reading, vocabulary and other skill areas for children below the tenth percentile. Thomas and Pelavin (1977)

report that the publishers of the Metropolitan Achievement Test found that the interpolated grade equivalent norms overestimated the actual reading gains of students during the school year. They also report that the norms for the Stanford Achievement Test were found to underestimate the actual empirical achievement of students. In other words, 1.0 month gains are harder to achieve from fall to spring on the Metropolitan Achievement Test and easier to obtain on the Stanford Achievement Test. It should be said that these results are no reflection on the quality of these tests. In fact, the publishers should be complimented for empirically testing the validity of the existing growth norms. The validity of interpolated norms on achievement tests are suspect because no data of any kind has been reported.

Pelavin and David (1977), and Stenner et al (1977) reported summer losses in terms of the percentile rank of the average spring and fall test scores. The percentile rank has certain disadvantages if the summer losses of different groups are being compared. As with grade equivalent scores, the absolute difference in achievement represented by the difference in percentile ranks changes from point to point in the distribution of scores. Horst, Tallmadge and Wood (1977) proposed the use of the normal curve equivalent (NCE) as a method of converting the percentile rank differences into a common metric. The NCE is a standard score with a mean of 50 and a standard deviation of 21.06. A unique NCE exists for each percentile rank and differences between NCE's are linear at all points in the normal distribution.

The use of NCE's to measure changes in achievement is based upon a number of untested assumptions about the characteristics of the norms of the standardized tests. Tallmadge and Wood (1976) suggest the following rules for using NCE's to measure achievement gain: 1) use a norm referenced test with an expanded standard score scale, 2) whenever possible, use the same level and form of the test for both the pretest and postest, 3) select a level of the test where the average child will get from 35 to 50 percent of the items correct, 4) perform all testing within two weeks of the empirical norm points, 5) use linear interpolation if the testing is within six weeks of the norm data, 6) choose all project participants on some test other than the pretest. With the Metropolitan Achievement Test it is often necessary to cross levels between the spring and fall testing. DeVito and Long (1977) and Pelavin and Barker (1976) report substantially larger summer losses when the test levels are not the same. In particular, Pelavin and Barker (1976) conclude that the levels of the Metropolitan Achievement Tests are not interchangeable. Slinde and Linn (1973) found "non-trivial" differences between different levels of several major tests based on their investigation of anchor test data. Stenner et al (1978) suggested that the levels of some tests are poorly articulated in terms of test content, and that the higher levels tend to be less representative of the expanding domain of instructional concepts and objectives.

The adequacy of the empirical test norms for measuring achievement gains has been questioned on other grounds. Tallmadge and Wood (1976) stated that the method is most appropriate when children in the treatment group are similar to those in the normative sample. It is possible that groups differing in socioeconomic status actually learn at rates quite different from those implied by the national norms.

Linn (1978) reported NCE losses in reading over three different standardized tests in eight cities. He also stated that the equipercentile norms are consistently overestimates of actual achievement growth for children in the low percentile ranks, especially if the children are black.

Other Technical Problems Associated with Achievement Growth. Bloom (1964) reviewed the literature and found that there had been very few long term longitudinal studies of educational achievement. Pelavin and David (1977) found that despite the massive testing associated with the Title I program evaluation, little data was available beyond the achievement gains made during the fall to spring period. The absence of long term empirical data on the achievement growth of the national population, or of sub-populations such as Title I project children, complicates the interpretation of longitudinal results.

The achievement growth estimates in the norm tables of the standardized tests were obtained on different groups of children measured during the same year. No existing test has norms based on the measured growth of the same children at several points in time. Unpublished evidence indicates that the standard score scales of

standardized tests tend to overestimate the longitudinal achievement growth of low scoring children, and underestimate the longitudinal achievement growth of high scoring children.

Stenner et al (1977) introduced the Standardized Growth Expectation (SGE) as a measure of the consistency of a standardized test's underlying achievement growth rate. The SGE is defined as "the amount of growth in achievement (expressed in standard deviation form) that a student must demonstrate over a given treatment interval to maintain his/her relative standing in the norm group". The difference between the pretest percentile and the postest percentile expressed in standard score form is termed the SGE. All five major standardized test batteries showed large decreases in SGE over the primary and intermediate grades. The decline of the SGE over time has been interpreted in the past as a typical learning curve. The assumption was made that achievement growth slows with maturation in the same way that physical growth slows. In his comprehensive survey of achievement growth, Bloom (1964) found that approximately fifty percent of general achievement at grade 12 (age 18) was reached by the end of grade 3 (age 9).

An example of the impact that the decline in SGE has on the longitudinal analysis of achievement gains can be made by considering the following example. A seven NCE (one-third of a standard deviation) increase in reading achievement at the second grade on the Metropolitan Achievement Test represents a 33 percent increase in learning above the expectation for second graders. This same seven NCE represents a 100 percent increase above

expectation for third graders, and a 200 percent increase above expectation for eighth graders. Stenner et al (1977) concludes that "any cross grade comparisons of treatment effects expressed as national percentiles, standard scores, NCE's or grade equivalents are usually inappropriate." They also state that upper grade level programs are likely to seem ineffective because the tests are no longer measuring anything affected by learning in school.

The use of the SGE to describe test characteristics revealed surprising anomalies in the amount of achievement growth that different tests expect over the summer months (see Attachment A). The Stanford Achievement Test norms assume that about ten percent of the annual reading achievement gain is made over the summer, while the Metropolitan Achievement Test norms assume that fifty percent or more of the annual reading achievement gain is made over the summer. Stenner et al (1977) suggests that tests with large summer SGE's are relatively insensitive to instruction.

A review of the literature on longitudinal growth in achievement on standardized tests raises more questions than it supplies firm answers. Describing school year gains in reading achievement and concomitant summer losses is much easier than interpreting them.

Stenner et al (1977) pointed out the importance of discovering the causes for summer loss before important policy decisions are made to correct it. Stenner et al (1978) suggests that the existence of summer loss may depend more on the particular norm-referenced test used to collect the data than it does on the actual learning of children in a particular program.

Methodology

The student achievement data included in this report were collected as part of the evaluation of the Growth in Academic Performance component of the ESEA Title I project in Cincinnati. Since September of 1975, the achievement growth of each child in the project has been maintained as part of the Title I Data Bank. The data bank contains fall and spring test data on over 7,000 children who have been identified as eligible for Title I classes during the 1975-76, 1976-77 and 1977-78 school years. The data bank was updated twice during each school year, with information on the mobility of the children within the Cincinnati public school system.

Data now exist on the two year reading achievement gains of two groups of Title I children from fall 1975 through fall 1977. The first group includes all children in the second grade during the 1975-76 school year who remained in the project during the third grade (1976-77) and who were pretested in the fourth grade during the current school year. The second group includes all children in the third grade during the 1975-76 school year who remained in the project during the fourth grade and who were pretested in the fifth grade during the current school year. Reading achievement scores are available at five consecutive points for all children in these two groups.

Changes in the Title I evaluation requirements and changes in local evaluation procedures produced changes in the test levels, forms and test dates used in the evaluation since 1975 (see Attachment B). Out-of-level testing (using a test form easier than

recommended by the publisher) was eliminated after the 1975-76 school year to permit calculation of NCE's, and to allow the use of the city-wide achievement test levels as the Title I postest. It was felt that the double testing of children each April was technically unsound and unnecessary. Pretest dates were moved up from the fourth week of September in 1975-76 to the second week of September 1976-77 to expedite the early selection of students for Title I classes. In 1977-78 children were selected for classes on the basis of the 1976-77 postest and the date of the pretest was established at the national norm date to meet the full requirements of the new Title I evaluation procedures. The postest in 1976-77 was approximately five weeks late because of the teacher strike.

Since many of the achievement comparisons cross either different forms or levels of the test, all of the comparisons of absolute changes in reading achievement will be made in terms of the standard score scale of the Metropolitan Achievement Test. Changes in the relative achievement of children across the various testing periods will be made using both percentile ranks and NCE's.

The consistency of summer losses (or gains) in reading achievement was tested by obtaining correlation coefficients between the changes in spring-to-fall achievement over the two summer periods available on each child. These correlations were conducted on the Normal Curve Equivalent changes over the two summer periods.

In cases where the tests were administered more than two weeks away from the empirical norm dates, interpolated corrections

were calculated for the obtained standard scores. These corrections were weighted on the basis of the distributions of obtained scores. The corrected standard scores, percentile ranks and NCE's are reported along with the obtained scores.

Results

Matched longitudinal scores were available on 295 children in the first group (2-3-4) and on 175 children in the second group (3-4-5). The original groups tested in both grades two and three in September 1975 contained approximately 700 children. The declines in both groups are due principally to the mobility of the Title I population and to the smaller proportion of Title I schools offering classes in the intermediate grades.

The potentially biasing effects of this level of sample mortality were estimated at each point by comparing the mean achievement levels of the matched group with the mean achievement levels of the total evaluation population at each testing point.

Question 1. Do Title I students show absolute losses in reading achievement over the summer period?

The mean standard scores in total reading for the matched group of each of the five testing points are given in Table 1. Group One showed an absolute summer loss of one standard score in 1976 and a gain of one standard score during the summer of 1977. Group Two showed a large absolute loss of six standard score units in the summer of 1976 and a moderate gain of three standard score units in the summer of 1977.

Table 1. Mean Standard Scores for Group One and Group Two at Each of the Five Testing Points.

	Testing Points				
	2 Fall	2 Spring Summer	3 Fall	3 Spring Summer	4 Fall
GROUP ONE (N=295)					
Observed Mean SS	32	43	42	49	50
Corrected Mean SS	32	43	44	48	50
Predicted Mean SS*	32	41	44	47	49
Observed Total Group SS**	31	44	42	50	

	3 Fall	3 Spring	4 Fall	4 Spring	5 Fall
GROUP TWO (n=175)					
Observed Mean SS	42	51	45	53	56
Corrected Mean SS	42	51	46	51	56
Predicted Mean SS*	42	44	47	51	54
Observed Total Group SS**	43	53	45	54	

*The equipercentile standard score predictions based on the initial mean standard score of each group.

**The mean standard score of all children pre and postested during each school year.

When the mean standard score units are corrected for testing dates, summer gains in standard scores are found in three of the four summer periods. The large summer loss recorded by Group Two between the third and fourth grades during 1976 was reduced slightly to five standard scores by the corrections.

The difference in size of the corrections at each test point are related to the rate of achievement gain in the empirical norms of the Metropolitan Achievement Test at the initial reading achievement level of each group. A predicted mean standard score is provided for each group indicating the pattern of standard score gains found in the norms for the fall 1975 mean scores of Group One and Group Two. Over the two year period, Group One gained one standard score above the prediction and Group Two gained two standard scores above the prediction.

The pre and post standard scores of the entire evaluation group of each group in both 1975-76 and 1976-77 are presented for comparison. In each case the means represent six to seven hundred children included in the Title I evaluation. The average scores of Group One and Group Two are very similar to the average scores of the total Title I population in each year. At no point is the difference more than two standard score units. There is a slight tendency for the matched group to show lower gains than those recorded for the total group. This would be expected because children showing large gains are systematically eliminated from the project if they score above the Title I eligibility cutoff. There is no evidence of a serious difference between the matched groups and the total population in the Title I project in reading achievement level.

Absolute summer loss in reading achievement was found in only one of the four periods included in the data. Group Two lost five standard score units during the summer between the third and fourth grade. Group One, however, gained two standard score units over the same grade period. The levels of the achievement test used to measure spring-fall achievement were different over this period for the two groups. The large summer loss in achievement in 1976 was measured in spring with the Primary II level, Form G, and in the fall with the Elementary level, Form F. The modest gain in achievement across the summer in 1977 was measured in spring with the Elementary level, Form G, and in the fall with the Elementary level, Form F.

Question 2. Do Title I students show relative losses in reading achievement over the summer period when compared to the national norm group?

The percentile ranks and normal curve equivalents showing the relative changes in reading achievement were based on the corrected standard scores (see Table 2). In Group One a summer loss of two NCE's was recorded during 1976 and no change was observed over the summer of 1977. In Group Two a fifteen NCE summer loss was recorded in 1976 and a four NCE summer loss was recorded in 1977.

Over the two year period both Group One and Group Two made small gains against the national norm group. Group One made a net two NCE achievement gain above prediction and Group Two made a four NCE gain above prediction. These small gains represent the combination of relatively large increases during the fall-spring period, followed by substantial losses during the summer period.

Table 2. The Percentile Ranks and NCE's Corresponding to Mean Standard Scores in Group One and Group Two, at each of the Five Testing Points.

	Testing Points				
	1975/1976		1976/1977		1977
	2 Fall	2 Spring	3 Fall	3 Spring	4 Fall
GROUP ONE (N=295)					
Percentile Rank*	10	12	10	12	12
Normal Curve Equivalent*	23	25	23	25	25

	3 Fall	3 Spring	4 Fall	4 Spring	5 Fall
GROUP TWO (N=175)					
Percentile Rank*	7	20	6	14	10
Normal Curve Equivalent*	19	32	17	27	23

*All norm based scores are based on the corrected standard score means.

The pattern of achievement gains during the third grade school year and summer are particularly different for the two groups. Group Two, pre and postested with the Primary II level of the Metropolitan, recorded a twelve NCE gain during the school year and then a fifteen NCE loss to the fall administration of the Elementary level. Group One recorded a two NCE gain during the school year with the Primary II level as the pretest and the Elementary level as the postest, but maintained its achieved level to the fall administration of the Elementary level. If the two school years were comparable, it would be safe to conclude that these two levels of the test are poorly articulated either in terms of test equation or content. The 1977 school year, however, was affected by the teacher strike and the impact of the severe winter. It is impossible to separate the effects of these events from the effects of changing test levels.

Question 3. Do Title I students show relative losses in reading achievement over the two year period when compared to other children in the Cincinnati public schools?

No data are available on the pattern of achievement gains for non-Title I schools during the summer period. The annual pattern of achievement gains of non-Title I students in Title I schools and non-Title I schools are estimated by comparing the achievement gains made by matched samples of students in all of the elementary schools between the April 1975 and April 1976 administrations of the city-wide achievement test. Matched achievement gains were obtained separately for Title I and non-Title I schools on a twenty percent sample of all children

tested at the two periods for grades two, three and four. The achievement gains were converted to NCE's so that the gains of groups with differing absolute achievement levels could be compared (see Table 3).

The results show that children in Title I and non-Title I schools consistently make annual reading gains smaller than those predicted by the national norms. This decline in reading achievement has been documented for years in the smaller proportions of elementary school children in Cincinnati elementary schools who score at or above the national norms in the higher grades. By making two assumptions the pattern of achievement growth of Title I children can be compared with the annual pattern of achievement growth shown by all children in the Title I and non-Title I schools. First, it must be assumed that the concurrent achievement growth rates determined between the 1975 and 1976 administrations of the city-wide achievement test can be used to estimate the true longitudinal growth of achievement over the period from end of first grade to end of fourth grade. This assumption would not be necessary if it had been possible to administer the city-wide achievement test during the spring of 1977. Second, it is necessary to assume that the annual achievement growth is linear for the period between the April administrations of the city-wide test.

Figure 1 illustrates the pattern of reading achievement growth made by Title I children in Group One compared with the national norm growth expectation and the estimated pattern of reading achievement growth made by a typical child in the

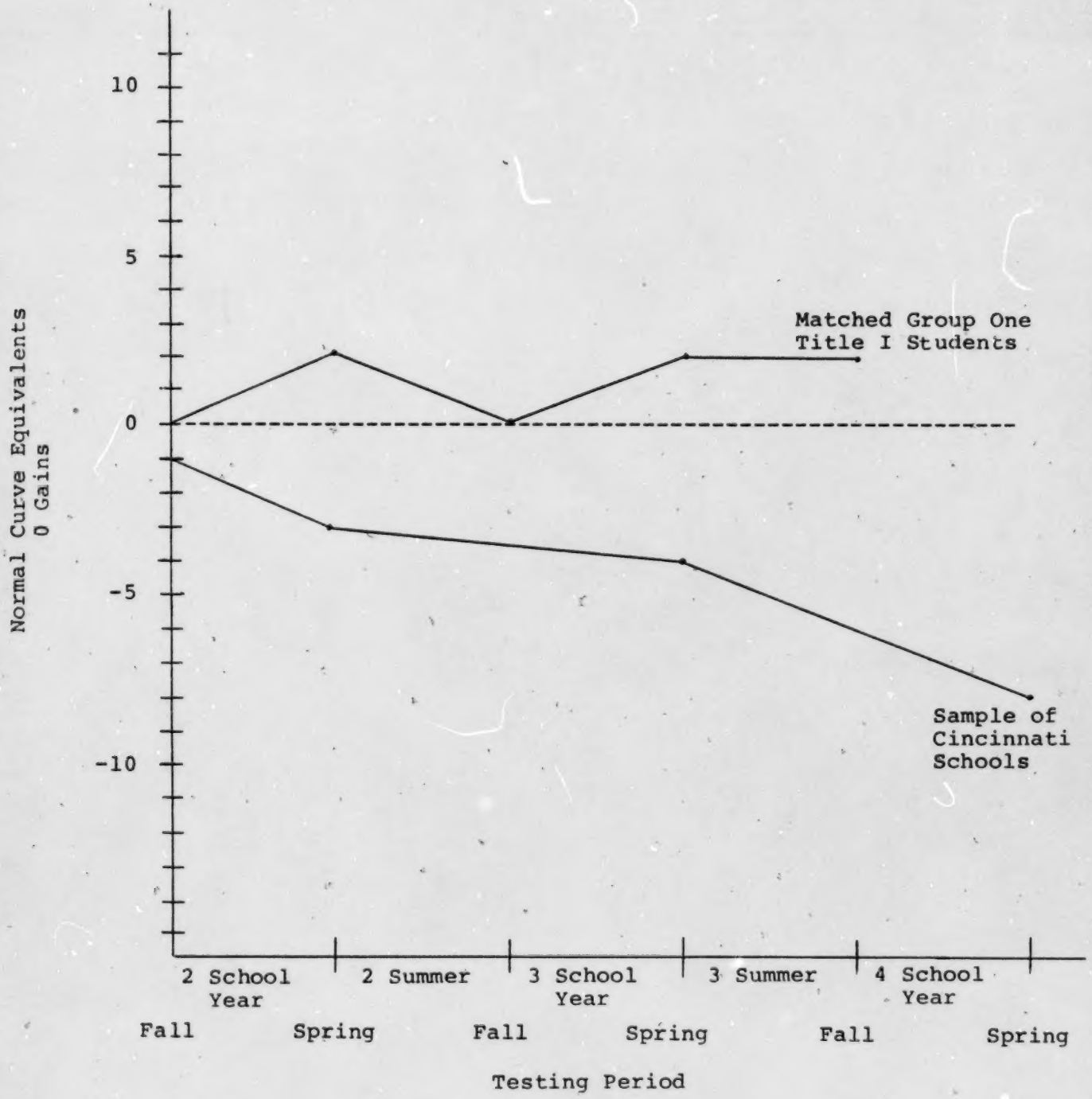
Table 3. Annual Reading Achievement Growth, in NCE's, Made in the Cincinnati Elementary Schools 1975-76 to 1976-77.

	Grade 1 to Grade 2		Grade 2 to Grade 3		Grade 3 to Grade 4	
	Spring	Spring	Spring	Spring	Spring	Spring
Sample in Title I Schools	-4 (N=294)		-1 (N=313)		-5 (N=345)	
Sample in Non-Title I Schools	-2 (N=397)		-1 (N=393)		-4 (N=403)	

The NCE Growth of Group One Compared With the Estimated NCE Growth of All Elementary School Children

Growth in Reading

FIGURE 1



elementary schools. The national norm expectation is shown by the horizontal line representing zero growth in NCE's. The expectation is that children will maintain their percentile rank, and hence show no growth on the NCE metric.

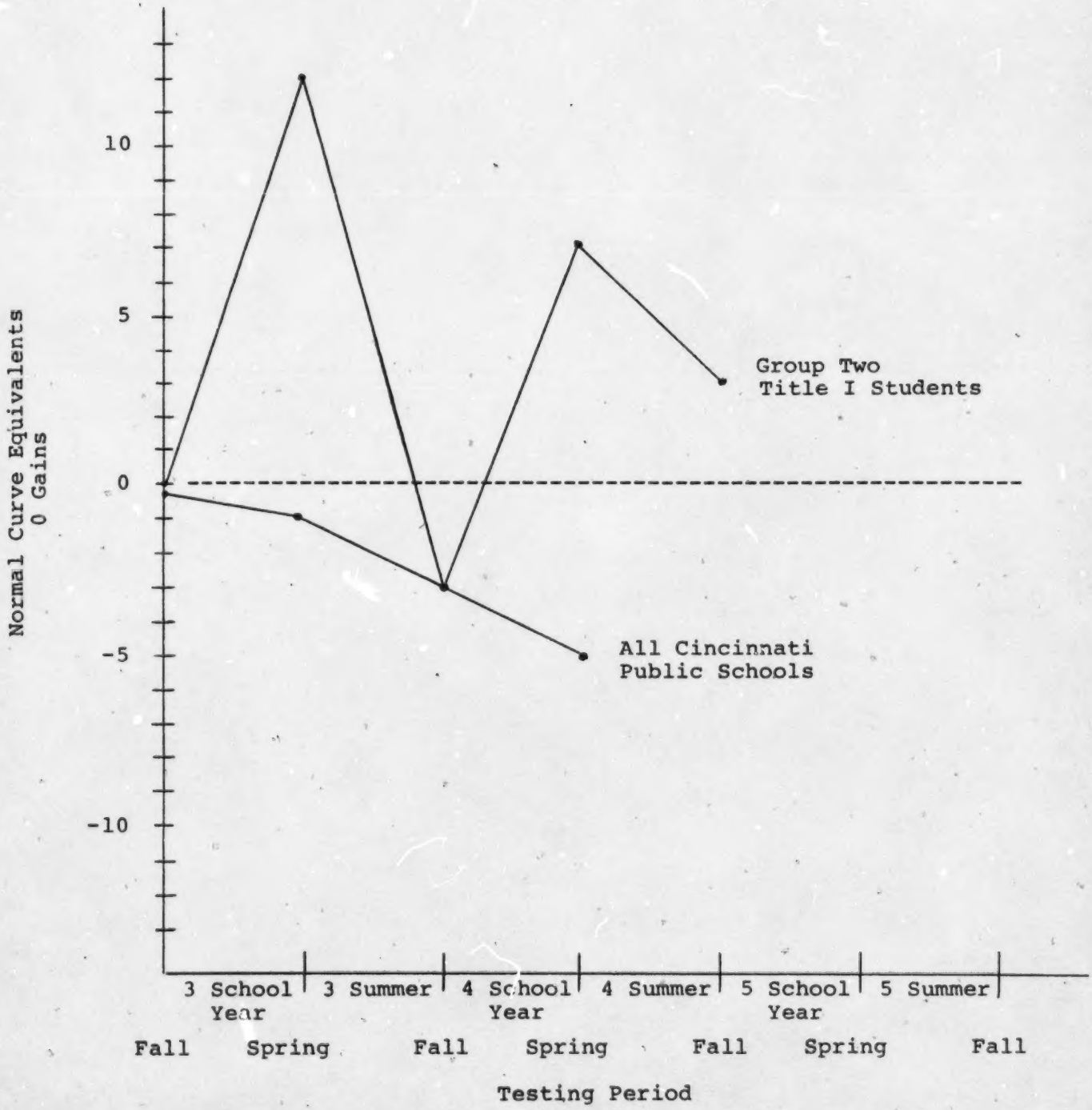
It can be seen from Figure 1 that after two years of Title I instruction the average child is two NCE's above predicted performance. During the same period, from the beginning of second grade until the beginning of fourth grade, the typical child in the elementary schools loses six NCE's. Comparing the Title I performance against the national norm shows a small cumulative improvement. A similar comparison against the estimated pattern of growth in the local elementary schools suggests a sizeable cumulative improvement. It is likely that this difference is exaggerated somewhat by the comparison of the performance of a matched longitudinal sample of Title I children with a matched concurrent sample of elementary school children. Rouse and Bollenbacher (1977) reported that the non-mobile school population in the Cincinnati elementary schools displayed better achievement than their classmates.

The performance of Group Two was compared with national and local performance in Figure 2. Since no fourth to fifth grade estimates of level achievement are available, the comparisons are possible only in the period from fall of the third grade to spring of the fourth grade. During this 18 month period, the Title I group shows evidence of large relative summer losses but an overall three NCE improvement against the national norm group. Over the same period, the estimated reading achievement of all

The NCE Growth of Group Two Compared With the Estimated NCE Growth of All Elementary School Children

Growth in Reading

FIGURE 2



elementary school children declines five NCE's below the predicted performance. The two year effects of the Title I program from grade three to grade five show a moderate three NCE gain versus the national norm, but an eight NCE gain versus the estimated performance of other children in the elementary schools.

Question 4. Is the pattern of achievement gain made by Title I children consistent over two successive summer periods?

It would be significant if it were determined that certain children, or perhaps schools, showed consistent patterns of summer loss across two consecutive summer periods. This would suggest that for certain children a summer program could be developed to sustain the school year achievement gains over the summer months. The NCE gains (or losses) for the two summer periods in both Group One and Group Two were correlated to determine the consistency, or predictability, of the summer losses over time. The NCE gains were also correlated with the total NCE gains made over the two-year period. High correlations in this latter case would indicate that the cumulative effects of the summer program on individual children are significantly related to their performance over the two year periods.

The results (see Table 4) show that there was little consistency in summer growth for individual children between the summer periods in either Group One or Group Two. Knowing the relative summer loss made by a child in 1976 would tell us virtually nothing about the likelihood of that loss being repeated in the summer of 1977. There was a modest relationship

Table 4. Correlations of Summer Loss 1976, Summer Loss 1977 and Total Gain in NCE's for Group One and Group Two.

	Summer Loss 1976 with Summer Loss 1977	Multiple Correlation of Summer 1976 and Summer 1977 with Total Gain
Group One (N=295)	.15	.38
Group Two (N=175)	.12	.41

in both groups between the cumulative effects of two periods of summer achievement and the total achievement gains made over the two year period. The observed correlations were not seriously affected by restriction of range. The standard deviations of the NCE gain over the summer periods showed a range of from 10 to 13, and the standard deviations of the total NCE gains over the two year periods were about 15 in both cases.

Summary and Conclusions

Given the number of factors which affect the measurement of patterns of achievement growth, and the relatively small amount of data on long term patterns of achievement growth, it would be easy to over-interpret the results reported in this study. The specific effects on absolute or relative achievement gains of crossing test levels and testing off the empirical norm dates are not known. The growing literature on the accuracy of empirical norms does, however, allow some evaluation of the impact of these procedures on the present results.

The generalizability of the results is also affected by the influence of the non-typical 1976-77 school year on the achievement of Title I children. These effects can be estimated in the future by comparison with the achievement gains on new groups of children in the Title I data bank. If the present results are to serve as an interpretable base for updated measurement of patterns of achievement growth, it is critical that changes in achievement test levels and achievement test batteries be studied by measuring the comparability of the new tests with the old ones. There is

sufficient evidence that achievement tests differ within a specific battery and across different tests enough to completely invalidate comparisons made on the basis of the published norms alone.

Question 1. Do Title I students show absolute losses in reading achievement over the summer period?

In most cases the level of achievement in fall is within one standard score of that measured during the previous spring. Data on the matched summer changes in achievement of all Title I children between spring 1976 and fall 1977 confirm the results obtained on the matched groups. The significant absolute summer loss of achievement between third and fourth grade in 1976 was not found in 1977. The large observed loss in 1976 was caused in part by the specific levels of the achievement test administered in spring and fall of 1976.

Content analysis of the specific questions on the Metropolitan Achievement Tests shows that the tests change drastically between the Primary II and Elementary levels. The Elementary level consists almost entirely of content area reading selections in science, social studies and health, where the Primary II level contains selections similar to basal reading texts. The Elementary form shifts to a relatively uncontrolled vocabulary level and places much more emphasis on inferential questions. The articulation of test content across grades should be considered seriously in the selection of any new achievement test for use in the city-wide testing program. Kasten and Tallmadge (1978) have documented studies which found important differences in the

vocabulary and reading skills measured between major reading tests. They also point out that certain tests have much higher content validity with specific basal reading series. The implications for measurement of the effects of any reading program are clear. Test selection should be made on the basis of assessed content validity of the instrument with respect to both the vocabulary and reading skills emphasized in the program.

Question 2. Do Title I students show relative losses in reading achievement over the summer period when compared to the national norm group?

Relative losses in reading achievement were found to occur at most grade levels over the spring-to-fall period. Some 80 to 90 percent of the achievement above prediction made during the fall to spring is lost during a two-year period during the spring to fall summer months. It is clear that at most grade levels the Title I students do not maintain the relative growth made during the school year over the summer months. These results must be interpreted carefully because the Metropolitan Achievement Test battery shows the highest level of expected summer gain of any of the five most used reading tests. Overall, half of reading achievement on the Metropolitan Test battery is observed to occur over the summer period. Stenner et al (1978) interpret this as either an indication that the Metropolitan Achievement Tests are relatively insensitive to the effects of school instruction, or that the norms are based on unrealistic assumptions made by the publishers in allocating annual achievement gain into the school year and summer periods.

The predicted standard score increases given in Table 1 illustrate the pattern of reading growth expected in the norm

tables. A typical child in Group One should gain nine standard scores during the second grade school year, three standard scores during the second to third grade summer, three standard scores during the third grade school year and two standard scores during the third to fourth grade summer. More than half of the absolute gain during the two years from the beginning of the second grade to the beginning of the fourth grade is expected to occur during the first six month period. A typical child in Group Two should gain two standard scores during the third grade school year, three standard scores during the summer between grade three and grade four, four standard scores during the fourth grade school year and three standard scores during the summer between grades four and five. In the period covered by Group Two, between the beginning of third grade and the beginning of fifth grade, exactly fifty percent of the reading achievement gains are expected to occur during the summer period.

It should be noted that the cumulative two year reading gains of Title I children in both Groups One and Two are positive. Compared to published data on the two year achievement gains of Title I children, it would appear that the Cincinnati Title I program is more successful than others in promoting sustained reading growth.

Question 3. Do Title I students show relative losses in reading achievement over the two year period when compared to other children in the Cincinnati public schools?

There is convincing evidence that children in the Title I program make better reading achievement gains than are made by

other children in the Cincinnati elementary schools. During their stay in the Title I program, Title I children move closer to the median reading achievement in the nation by a small amount, while the typical pattern of reading achievement gain in the elementary schools is away from the national norm level. The pattern of achievement growth in the city as a whole is consistent with the losses in reading achievement reported by Linn (1978) on eight major city school systems.

Question 4. Is the pattern of achievement gain made by Title I children consistent over two successive summer periods?

Relative summer loss in reading achievement is not predictable from year to year in the two groups analyzed in this report. There is no evidence that particular schools or children could be identified on the basis of summer loss for special programs to change the reading achievement pattern. The data provides some evidence that better achievement gains can be predicted on the basis of two consecutive measurements of summer achievement. Since the average child's stay in the Title I program is less than three years, it is unlikely that summer reading loss could be profitably used to select children for a summer reading program.

None of the evidence presented in this report suggests that an effective summer program would or would not be able to change the observed pattern of summer loss relative to the norm group. The evidence does confirm that relative summer losses in reading achievement are typical in the Title I program. A summer program which maintained the relative achievement gains made during the

school year over the summer months would greatly increase the cumulative impact of the Title I reading program. The effects of any proposed summer program should be assessed not over the summer period, but in relation to the annual and biennial pattern of achievement growth.

ATTACHMENT A
A Comparison of Summer and School Year
Achievement Gains on Five Achievement Tests
After Data Presented in Stenner et al (1978)

RATIO OF SUMMER GROWTH EXPECTATION TO SCHOOL YEAR EXPECTATION (50%TILE)

	TOTAL READING					TOTAL MATH				
Grade	Stanford	ITBS	CAT-77	CTBS	Metropolitan	Stanford	ITBS	CAT-77	CTBS	Metropolitan
2	.11	.51	.41	.70	.41	.16	.16	.29	.85	.30
3	.03	.53	.62	.56	1.00	.06	.66	.30	.42	.30
4	.04	.56	.65	.50	?	**	.64	.48	.93	?
5	.13	.66	.60	.42	.60	.16	.57	.40	.67	1.00
6	**	.61	.76	.52	1.00	.06	.61	.37	.65	1.67
7	.36	.54	1.13	.56	3.60	.50	.71	1.11	1.00	7.20
8	.11	.36	.86	?	1.00	.18	.50	1.14	?	1.00

*As an illustration the growth expectation over the summer interval on the CTBS Reading scale (for second graders) is 70% of the regular school year growth expectation for a 50th percentile student.

**Growth expectation over the summer period is zero.

?-Could not compute this value from information provided in the MAT Manual.

ATTACHMENT B

Test Level and Form Given by
Grade in the Title I Evaluation

	1975 Pre	1976 Post	1976 Pre	1977 Post	1978 Pre
Grade 2	Primary I 4th wk Sept Form G	Primary I 4th wk April Form G	Primary I 2nd wk Sept Form F	Primary II 3rd wk May Form G	Primary I Oct. 15* Form F
Grade 3	Primary II Form G	Primary II Form G	Primary II Form F	Elementary Form G	Primary II Form F
Grade 4	Elementary Form G	Elementary Form G	Elementary Form F	Elementary Form G	Elementary Form F
Grade 5	Elementary Form G	Elementary Form G	Intermediate Form F	Intermediate Form G	Intermediate Form F
Grade 6	Intermediate Form G	Intermediate Form G	Intermediate Form F	Intermediate Form G	Intermediate Form F

*Testing Date

References

Beggs, D.L. and Hieronymus, A.N. Uniformity of achievement growth, Journal of Educational Measurement, 1968, 5, 91-97.

Bloom, B.S. Stability and Change in Human Characteristics, John Wiley and Sons, New York, 1964.

DeVito, P.J. and Long, J.V. The effects of spring-spring vs. fall-spring testing upon the evaluation of compensatory education programs. A paper presented at the American Educational Research Association meeting in New York, 1977.

Hayes, D.P. and Grather, J. The school year and vacations: When do students learn? Washington, D.C., Urban Institute, 1969.

Hendrickson, G. Review of Title I evaluation studies. A paper presented at the American Education Research Association Convention, New York, 1977.

Holthouse, N.D. Stoffiet, F.P. and Tokor, E.B. Achievement, social class and the summer vacation. A paper presented at the American Educational Research Association Convention, New York, 1977.

Jaeger, R.M. The effect of test selection on Title I project impact. A paper presented at the American Educational Research Association Convention, Toronto, 1978.

Linn, R.L. The validity of inferences based on the proposed Title I evaluation models. A paper presented at the American Educational Research Association Convention, Toronto, 1978.

Ozenne, D. The effect of vertical scaling imprecision in the estimates of Title I project gains. A paper presented at the American Educational Research Association Convention, Toronto, 1978.

Pelavin, S.H. and Barker, P. A study of the generalizability of the results of a standardized achievement test. A paper presented at the American Educational Research Convention, San Francisco, 1976.

Pelavin, S.H. and David J. Evaluating long term achievement: an analysis of longitudinal data from compensatory education programs. Menlo Park, California: Project 4537, Stanford Research Institute, 1977.

Rouse, A.H. and Bollenbacher, J. A follow-up study of pupil achievement, Cincinnati Public Schools, 1977.

Scott, L.F. Summer loss in modern and traditional elementary school math programs, Journal of Educational Research, California, 1967, 18, 145-151.

Slinde, J.A. and Linn, R.L. Vertically equated tests: Fact or phantom? Journal of Educational Measurement, 1977, 14, 23-32.

Stenner, J.A. et al. Education of the ESEA Title I programs of the public schools of the District of Columbia, Washington, D.C., 1977.

Stenner, J.A., Hunter, E.L., Bland, J.D. and Cooper, M.L. The Standardized growth expectation implications for educational evaluation. A paper presented at the American Educational Research Association Convention, Toronto, 1978.

Tallmadge, G.K. and Wood, C.T. Users guide: ESEA Title I evaluation and reporting system, RMC Research Component, Mountain View, California, 1976.

Thomas, T.C. and Pelavin, S.H. Patterns in ESEA Title I reading achievement. Menlo Park, California: Project 4537, Stanford Research Institute, 1976.

Thurstone, L.L. A method of scaling psychological and educational tests, Journal of Education Psychology, 1925, 16, 433-451.

No Author. Summer recess: Does it make a difference on Title I student achievement. A paper presented at the American Educational Research Association Convention, New York, 1977.